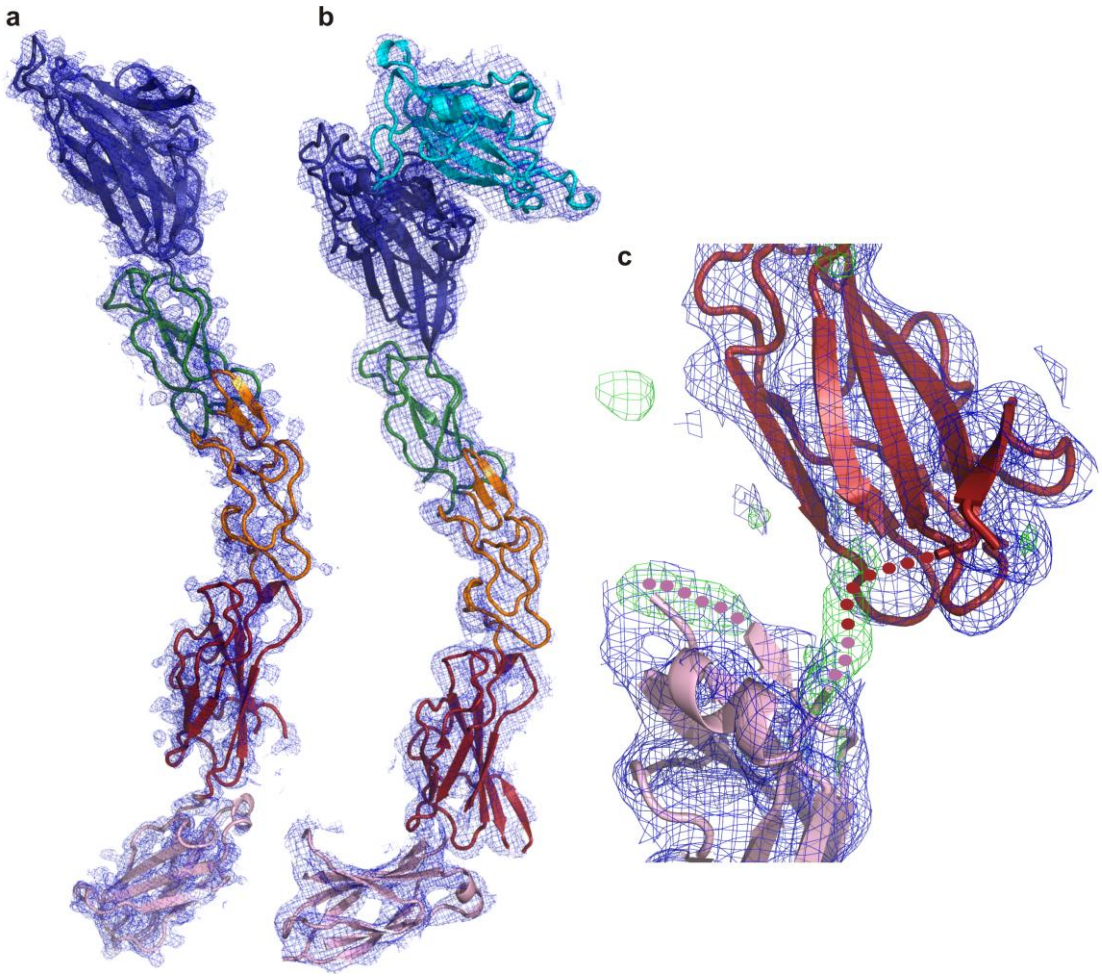
# An extracellular steric seeding mechanism for Eph-ephrin signaling platform assembly

Elena Seiradake[1], Karl Harlos[1], Geoff Sutton[1], A. Radu Aricescu[1] & E. Yvonne Jones[1]

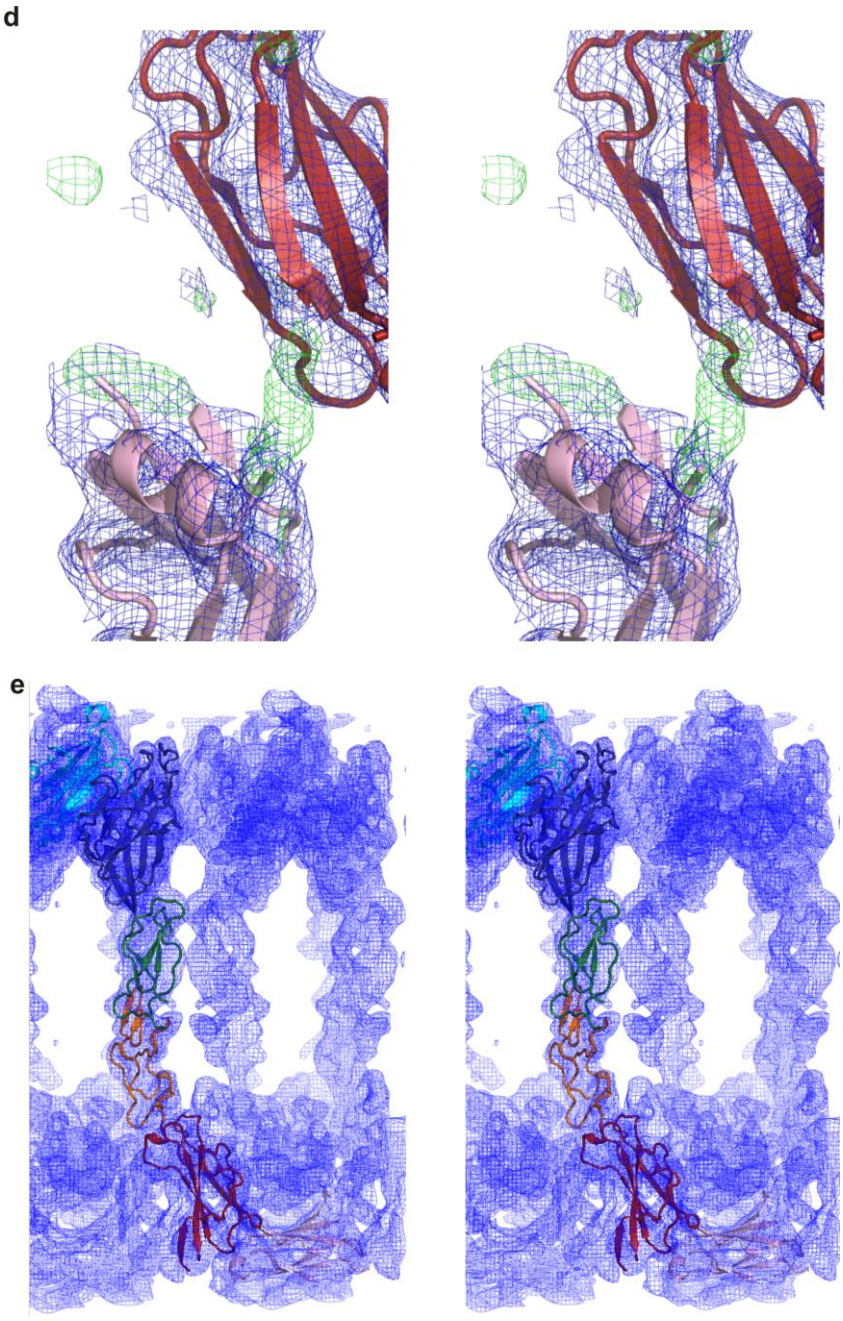
[1]*Division of Structural Biology, Wellcome Trust Centre for Human Genetics, University of Oxford, Roosevelt Drive, Oxford OX3 7BN, United Kingdom. Correspondence and requests for materials should be addressed to A.R.A (radu@strubi.ox.ac.uk) or E.Y.J. (yvonne@strubi.ox.ac.uk).*

# Supplementary Information

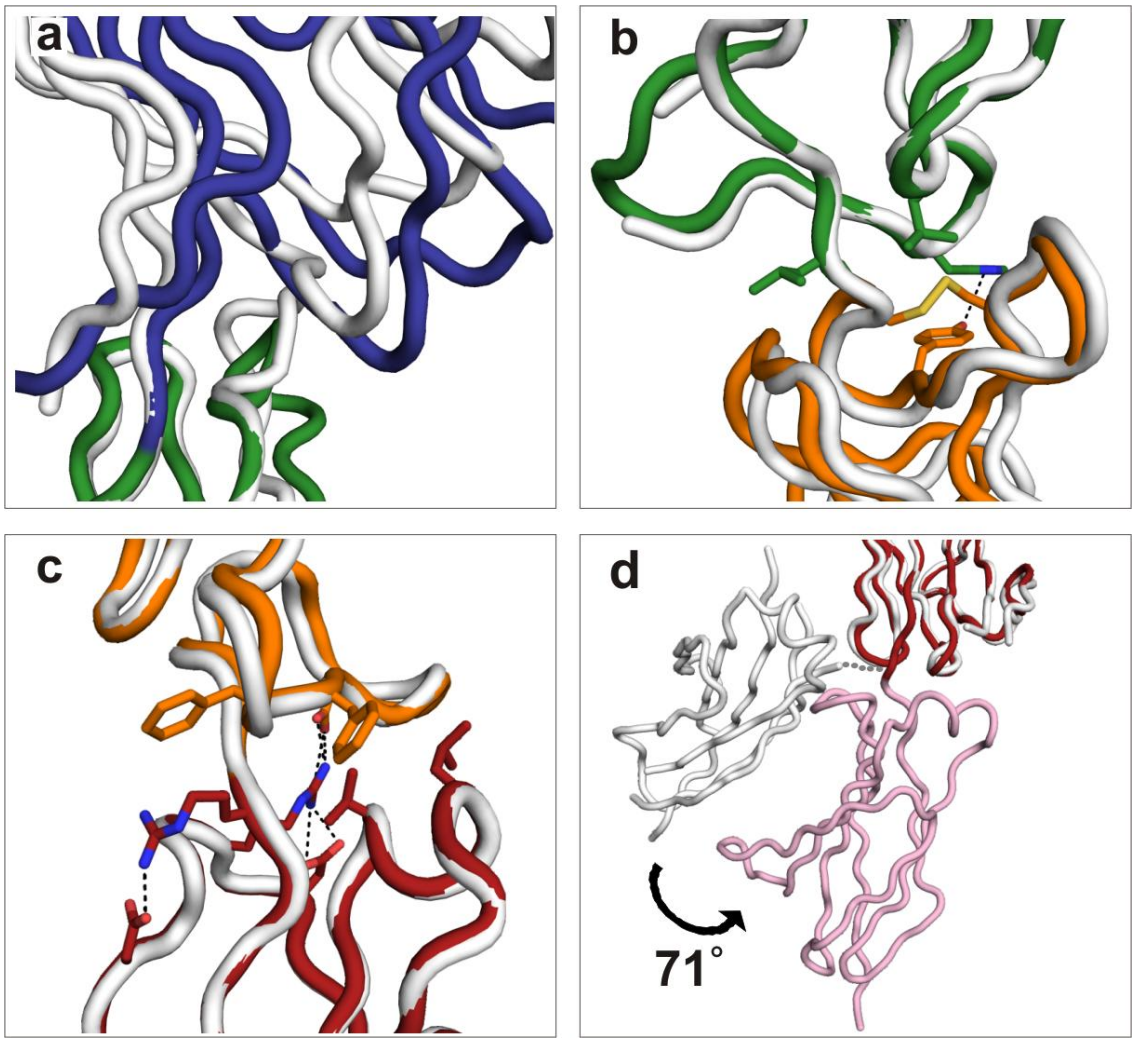
Nature Structural & Molecular Biology: doi:10.1038/nsmb.1782



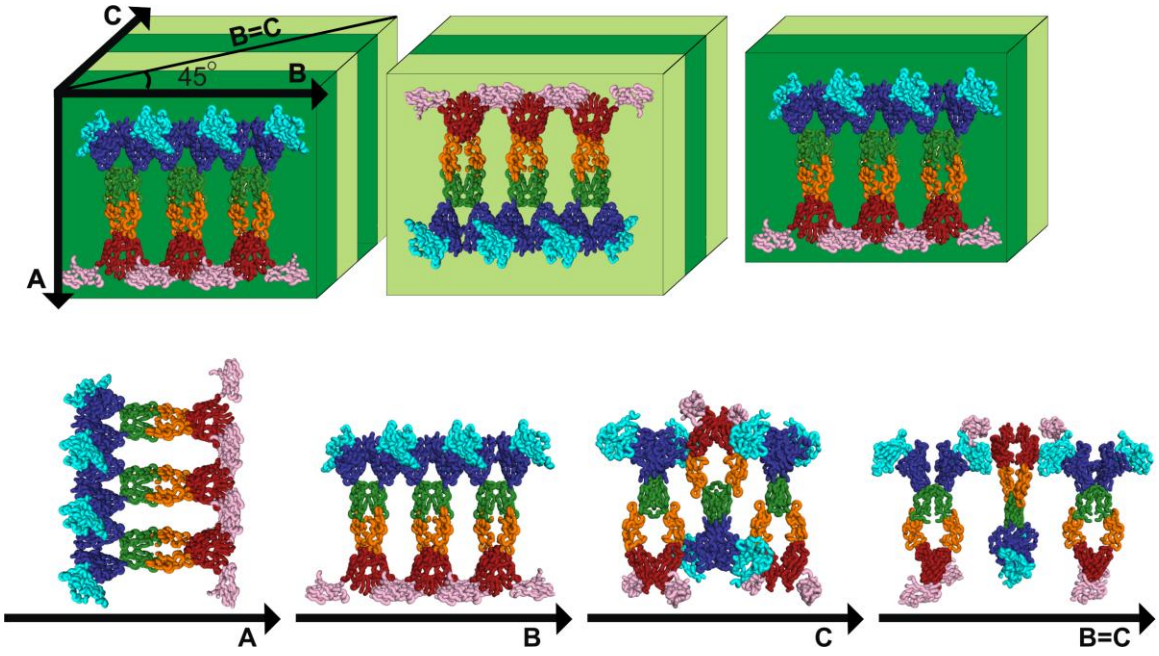

Nature Structural & Molecular Biology: doi:10.1038/nsmb.1782


d
e

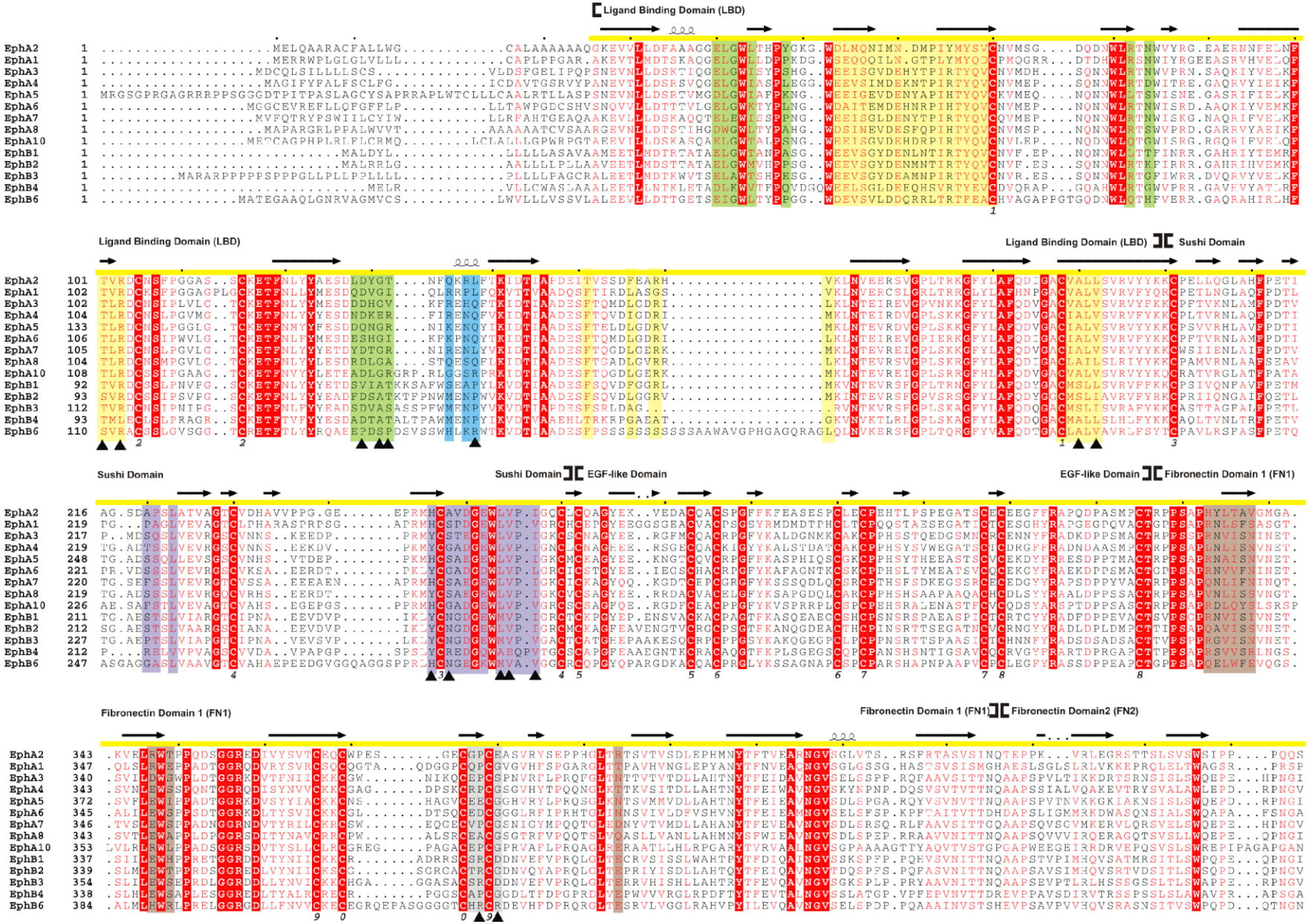
**Supplementary Fig. 1 | Structures of the full length extracellular region of EphA2 (eEphA2) alone and in complex with ephrinA5$^{RBD}$.** **a,** Crystal structure of eEphA2 (3 Å resolution). The colours of the ribbon diagram are LBD (blue), sushi domain (green), EGF-like domain (yellow), FN1 (red), FN2 (pink). The 2Fo-Fc electron density map (contour level = +1σ) is shown in blue. One molecule is displayed. Map contours were cropped at 3 Å distance from surface atoms in PYMOL (carve = 3 Å). **b,** Crystal structure of eEphA2-ephrinA5$^{RBD}$ (4.8 Å resolution). Colours and map settings are as in a, with ephrinA5$^{RBD}$ in cyan. **c,** Zoom on the FN1 and FN2 domains of eEphA2 in complex with ephrinA5$^{RBD}$. 2Fo-Fc (settings as in a) and Fo-Fc (contour level = +3σ, shown in green, carve = 10 Å) maps are displayed. The FN1-FN2 linker changes conformation compared to the unliganded eEphA2 structure used as search model and was not included in the lower resolution complex model. The loop comprising residues 513-515 (located within interface F) is not structured in the unliganded eEphA2 used as search model and was therefore not included. The putative positions of the missing chains are indicated as dotted lines. **d,** As in c, but a wall-eye stereo view is shown and dotted lines are omitted. **e,** Wall-eye stereo view of a larger excerpt from the 2Fc-Fo electron density map calculated from the complex crystal data after model refinement is shown in blue (contour level = +1σ). The carve command was not used, therefore all features of the electron density map are visible. One molecule of the eEphA2 and ephrinA5$^{RBD}$ complex is shown as ribbons.

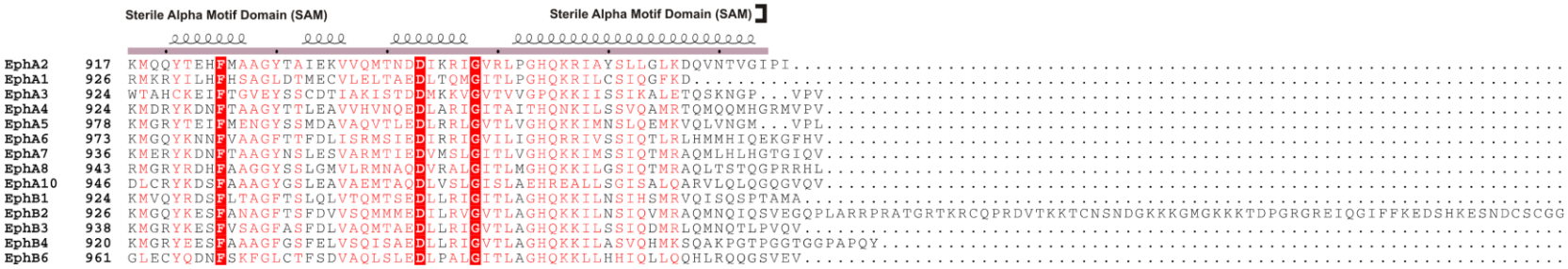



**Supplementary Fig. 2 | eEphA2 has flexible interdomain hinge regions between the fibronectin domains FN1 and FN2 and between the LBD and sushi domain.** Superposition of eEphA2 when in complex with ephrinA5$^{RBD}$ (white) and unliganded eEphA2 (colours are as in **Supplementary Fig. 1**: blue= LBD, green=sushi domain, orange=EGF-like domain, red=FN1, pink=FN2). Sidechains that provide rigidity to inter-domain regions are depicted as sticks. Hydrogen bonds are shown as intermittent black lines. **a,** LBD and sushi domains. The structures are superposed via the sushi domains only. **b,** Sushi and EGF-like domains. The structures are superposed via the sushi domains only. **c,** EGF-like and FN1 domains. The structures are superposed via the FN1 domains. **d,** FN1 and FN2 domains. The structures are superposed via the FN1 domains. The linker between FN1 and FN2 in the ephrinA5$^{RBD}$-bound structure is not modelled and indicated as a dotted grey line.

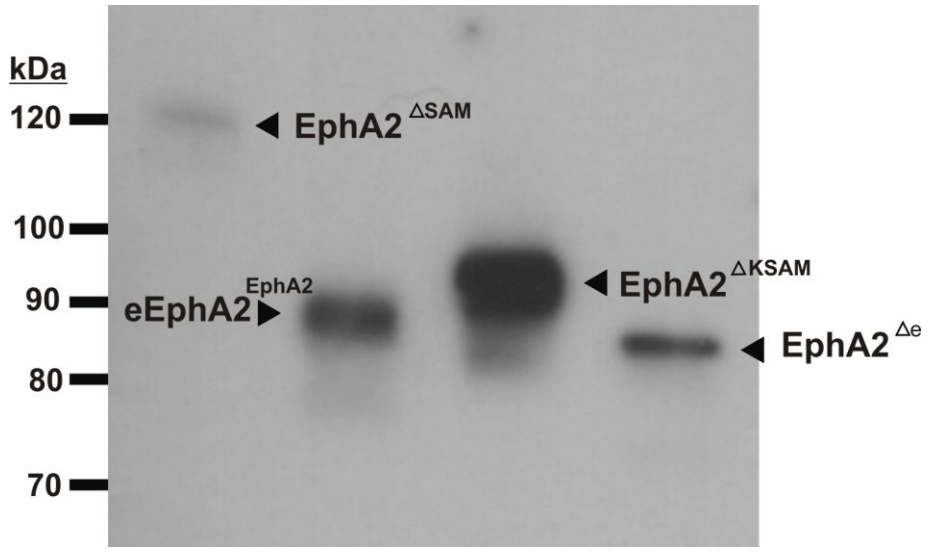



**Supplementary Fig. 3 | Crystal packing of eEphA2-ephrinA5$^{RBD}$ complex.** eEphA2-ephrinA5$^{RBD}$ complexes are shown as coil and coloured as in **Supplementary Fig. 1**. They pack into parallel arrays along unit cell axis B. These arrays stack on top of each other along unit cell axis A. The parallel arrays are packed in an anti-parallel fashion along unit cell axis C.





**Supplementary Fig. 4 | Sequence alignment of human Eph receptors.** The red shading of sequence indicates degree of conservation as calculated with ESPRIPT. Secondary structure elements are depicted above the alignment as are coloured bars indicating which example structure was chosen as template: yellow = extracellular region of EphA2 (eEphA2), green = transmembrane region of EphA1 (Protein Data Base code 2K1L[1]), blue = kinase domain and juxtamembrane region of EphB2 (Protein Data Base code 1JPA[2]), violet = SAM domain of EphA4 (Protein Data Base code 1BOX[3]). The residues corresponding to those within eEphA2-ephrinA5$^{RBD}$ crystal interfaces A-F are shaded with coloured background: yellow = A, green = B, blue = C, violet = D, grey = E, brown = F. Arrows mark the position of residues in EphA2 that were mutated in this study.




**Supplementary Fig. 5 | Deletion constructs of EphA2 were successfully expressed at the cell surface.** EphA2 constructs (shown in **Supplementary Fig. 6**) were expressed in HEK293T. 24 hours post-transfection, the cell surface proteins were biotinylated and purified using streptavidin-conjugated agarose beads. The resulting protein samples were applied to a denaturing polyacrylamide gradient gel, transferred to a nitrocellulose membrane and analysed by immunoblotting using an antibody against the C-terminal histidine tag.

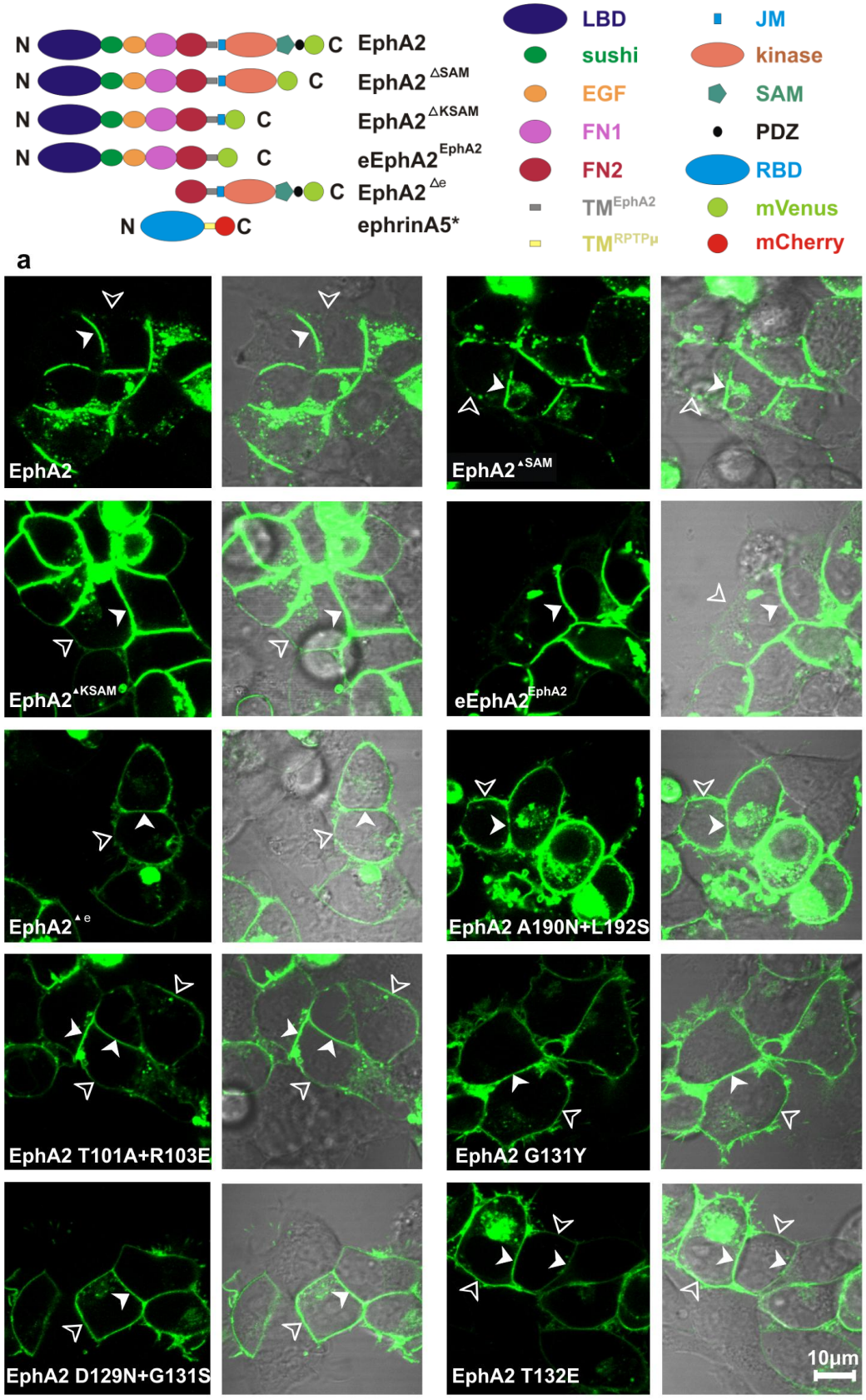

N
C
EphA2
EphA2 ΔSAM
EphA2 ΔKSAM
eEphA2 EphA2
EphA2 Δe
ephrinA5*
LBD
sushi
EGF
FN1
FN2
TM EphA2
TM RPTPμ
JM
kinase
SAM
PDZ
RBD
mVenus
mCherry
a
EphA2
EphA2 ΔSAM
EphA2 ΔKSAM
eEphA2 EphA2
EphA2 Δe
EphA2 A190N+L192S
EphA2 T101A+R103E
EphA2 G131Y
EphA2 D129N+G131S
EphA2 T132E
10μm

*Fig. 6a continued*

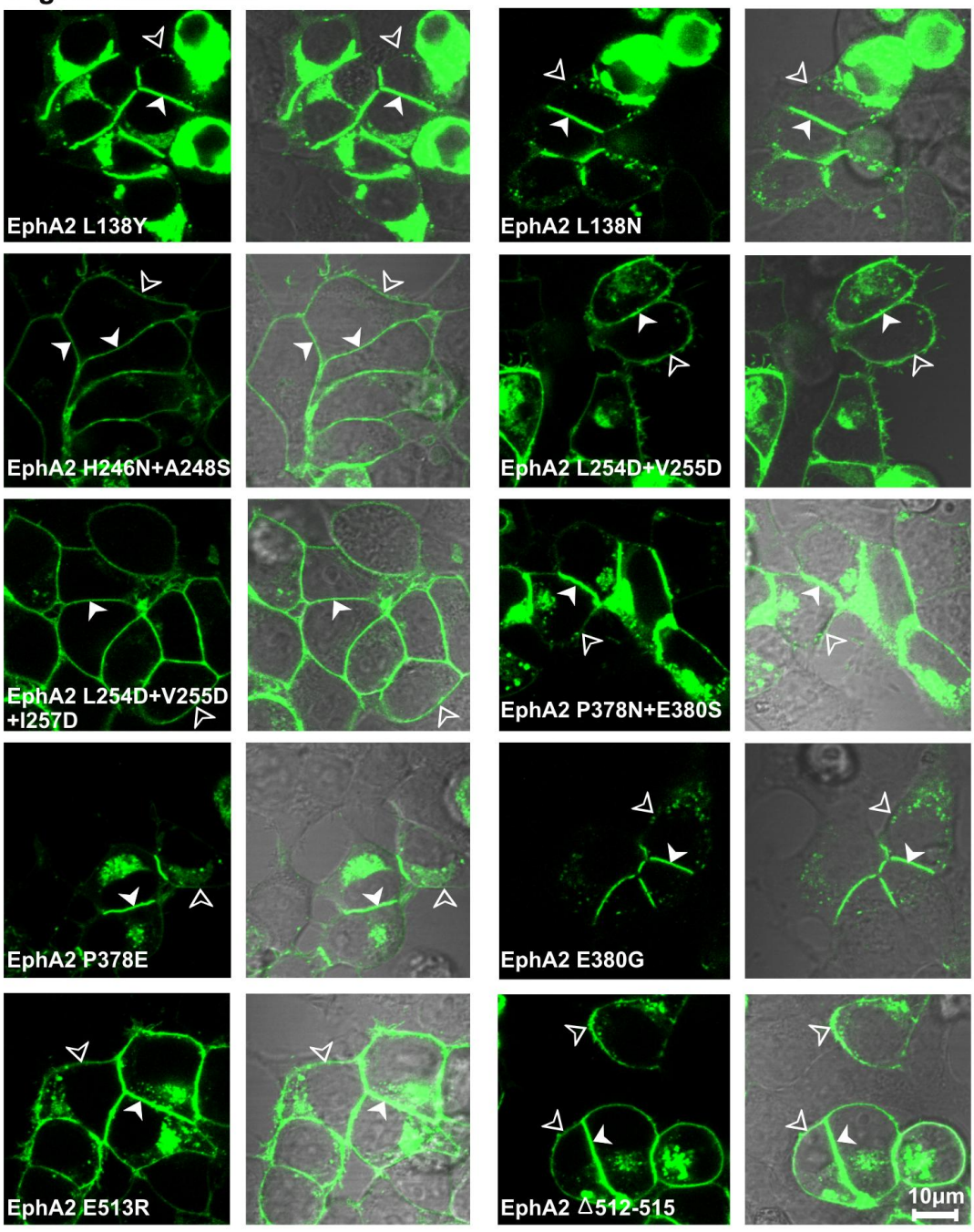
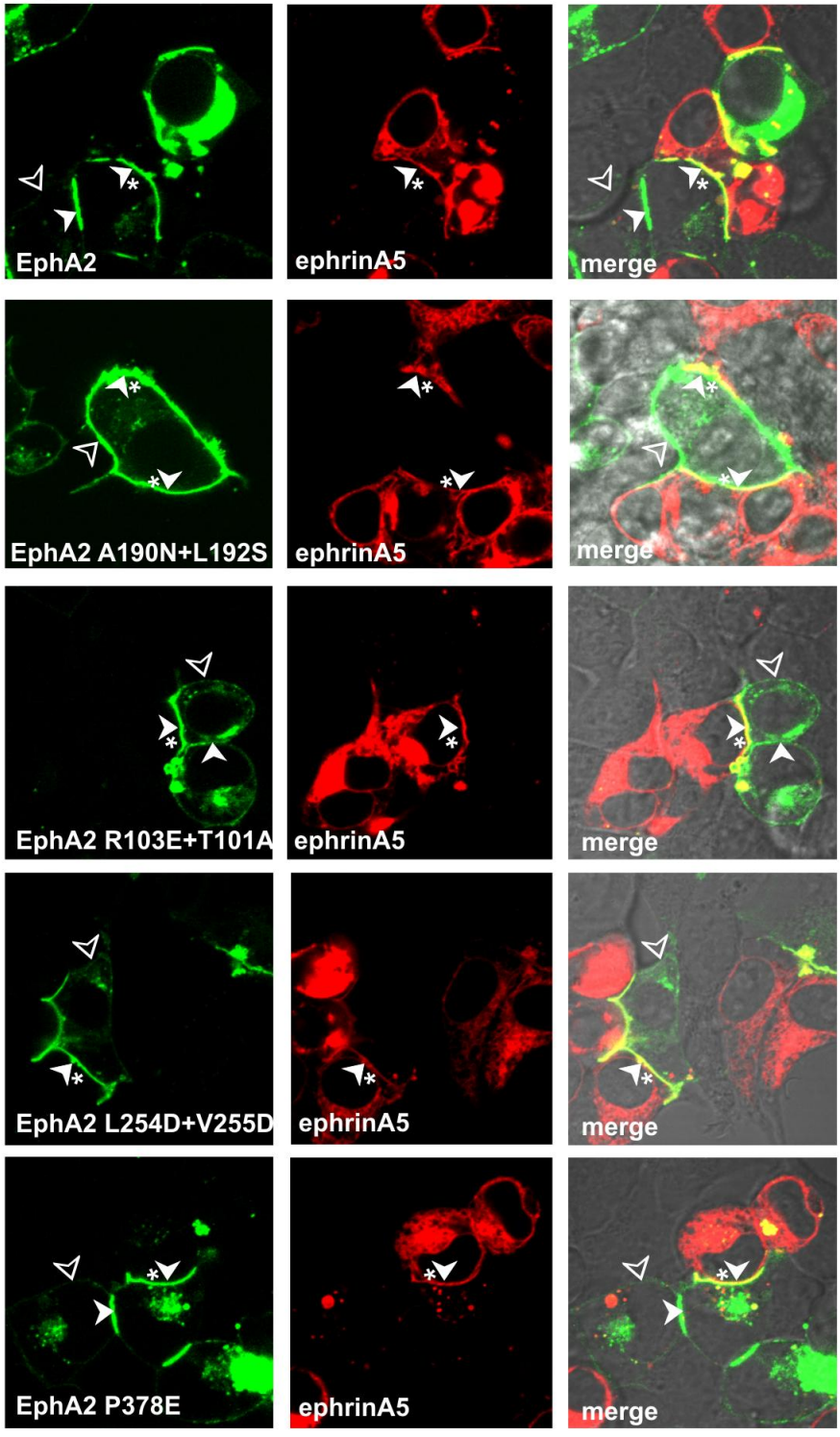


*Fig. 6a continued*



**b**

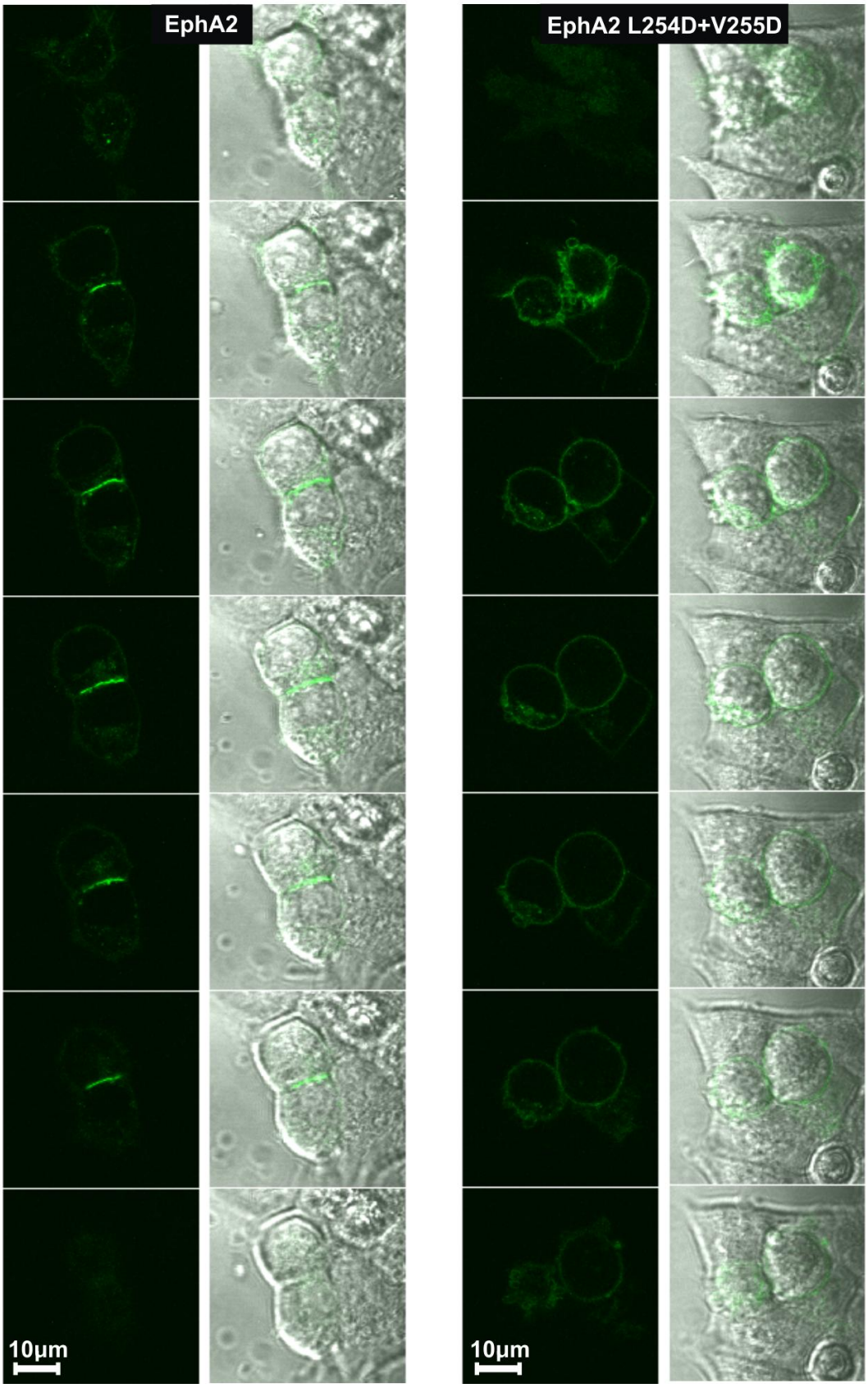
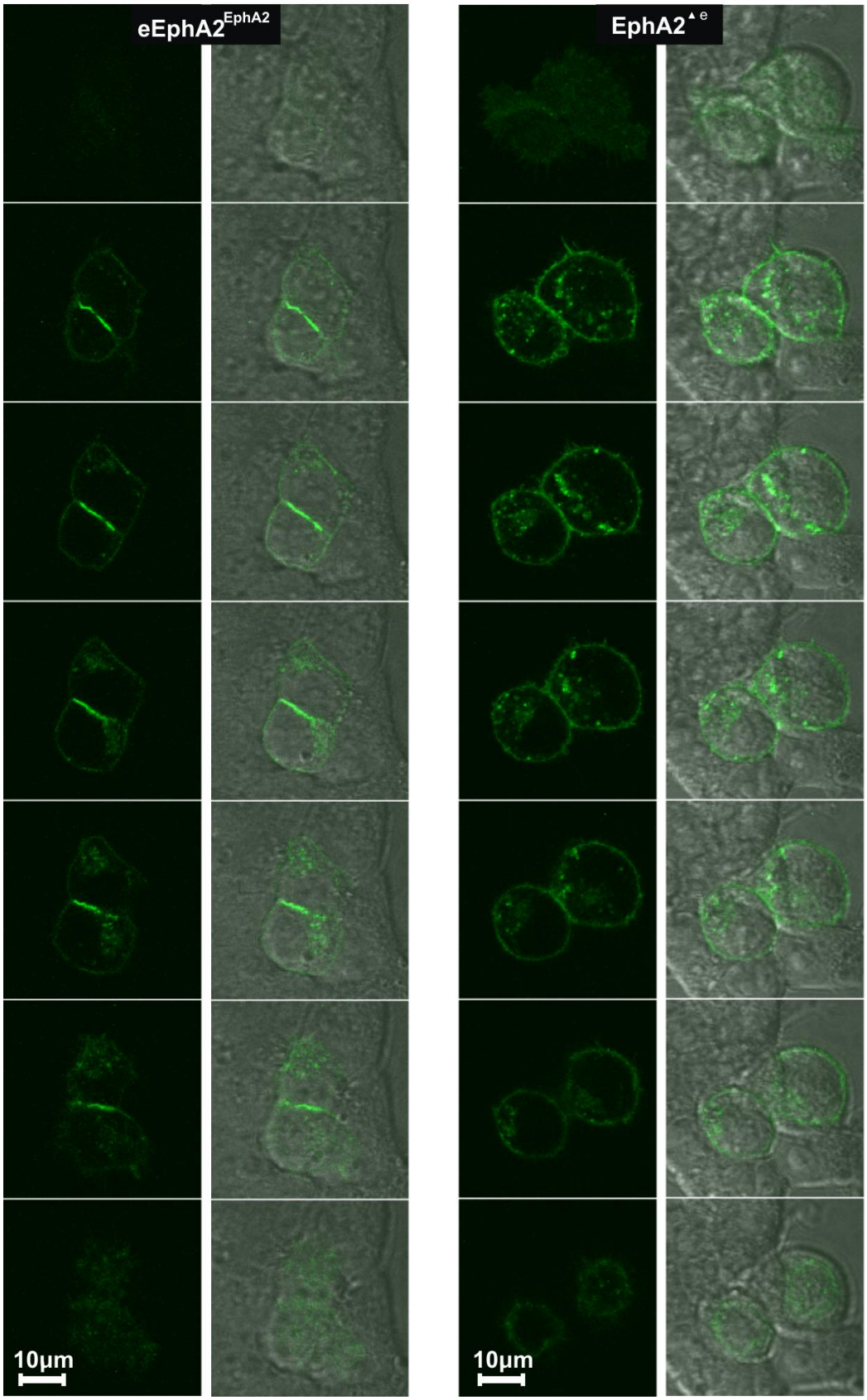
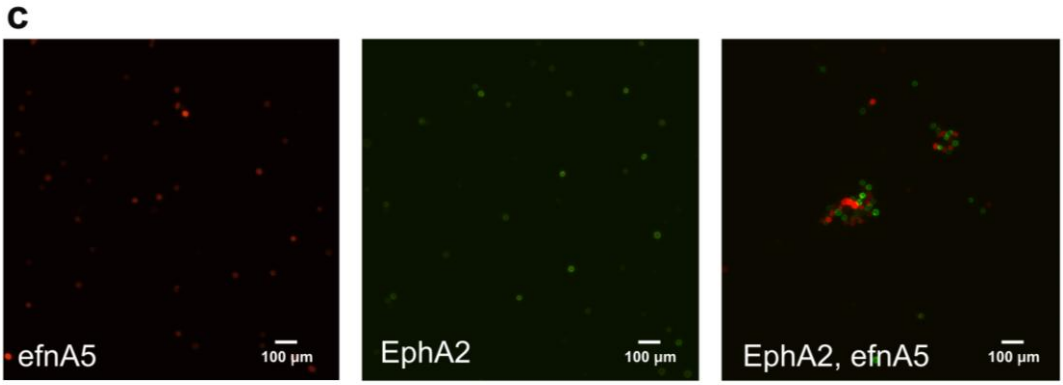
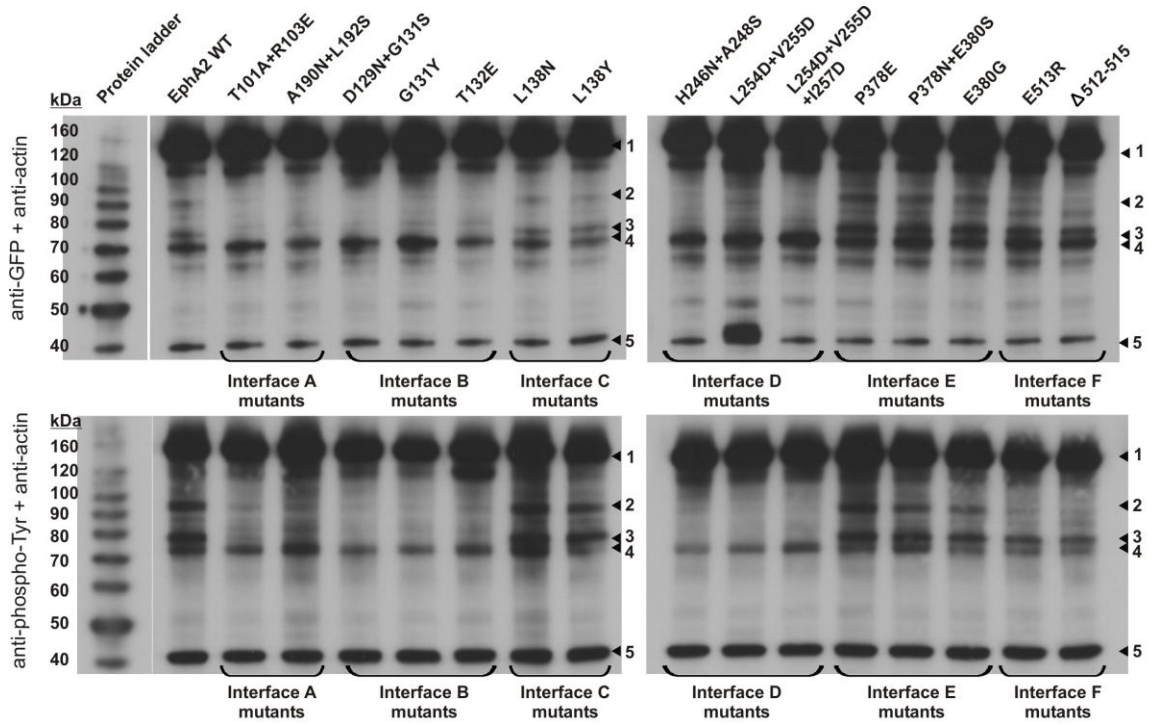


***Fig. 6b continued***

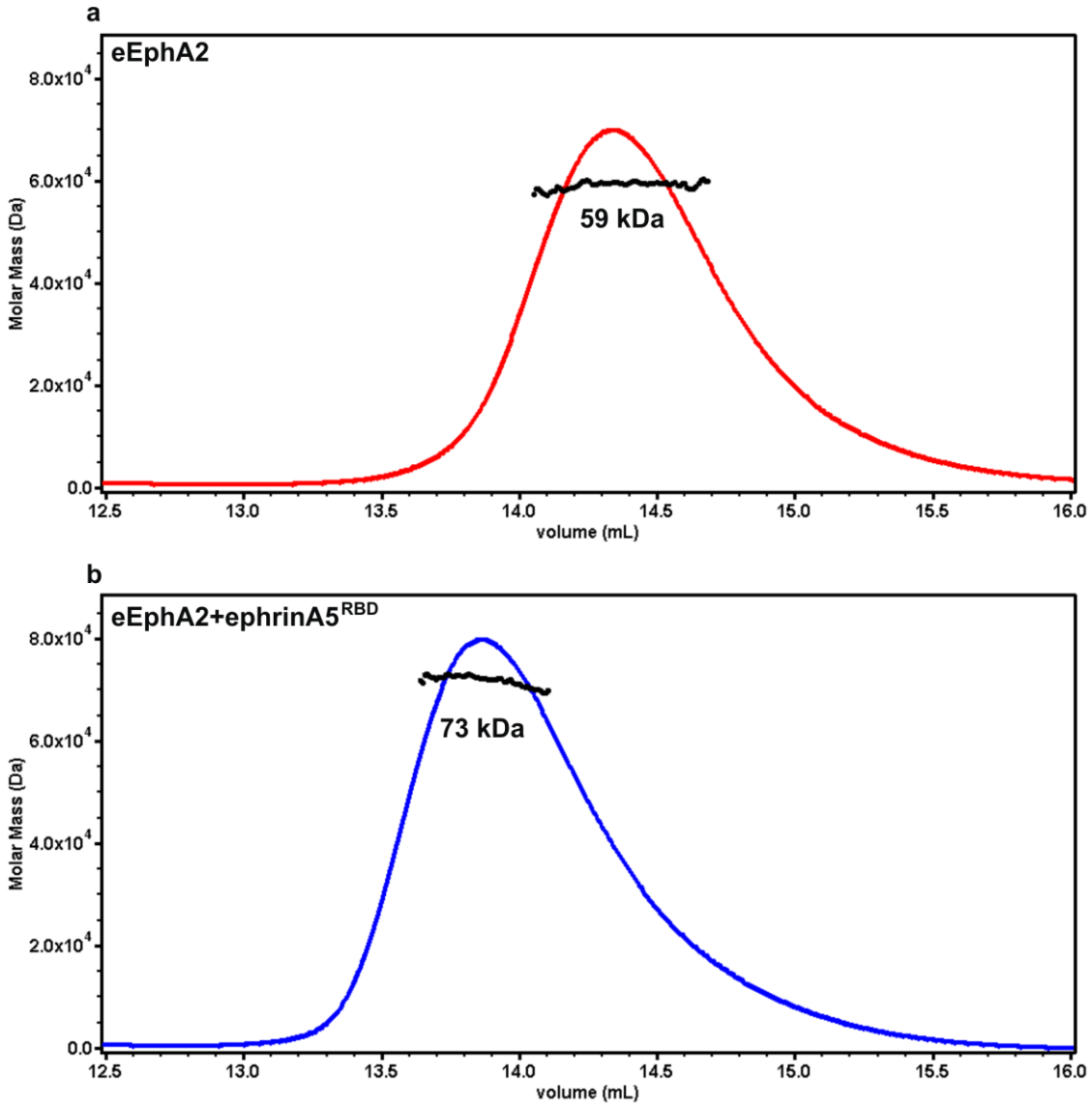





**Supplementary Fig. 6 | Expression of mVenus-tagged EphA2 and mCherry-tagged ephrin A5 constructs in HEK293T cells (a, b) and in insect cells (c). a,** EphA2 and ephrinA5$^{RBD}$ constructs are depicted schematically at the top of the figure. Filled white arrows indicate examples of cell-cell contact between cells that were both transfected with mVenus-tagged EphA2-derived construct. Open arrows point to contact between a transfected cell and a non-transfected cell or void. EphA2 point mutations are within the crystal contact interfaces A (A190N+L192S, T101A+R103E), B (G131Y, D129N+G131S, T132E), C (L138Y, L138N), D (H246N+A248S, L254D+V255D, L254D+V255D+I257D), E (P378N+E380S, P378E, E380G), F (E513R, Δ512-515). Note that EphA2 wild type and interface C and E mutants localize at cell-cell contacts of transfected cells, while mutants in interfaces A, B, D and, to some extent, F do not.
To observe *trans* interactions of cells expressing different EphA2 constructs and a membrane-tethered ephrinA5$^{RBD}$ construct, cells were transfected separately for 4 hours, re-suspended in fresh medium and plated out together. To tether ephrinA5$^{RBD}$ to the cellular membrane it was fused to the monomeric transmembrane helix of the receptor protein phosphatase RPTPμ. Filled white arrows with an asterisk point to contacts between cells carrying mVenus-tagged EphA2 constructs and cells carrying the mCherry-tagged construct ephrinA5. Note that EphA2 mutants in interfaces B-F localise at contacts with cells transfected with ephrinA5, suggesting that localization of cell surface EphA2 can be driven by ephrinA5 present *in trans*; presumably via the well-characterized ephrin binding interface (interface A). Introducing a glycosylation site in interface A (A190N+L192S) abolishes ephrinA5 binding in *trans*, while a double point mutant in interface A (EphA2 T101A+R103E) does accumulate at contacts with cells over-expressing ephrinA5, suggesting that it retains some ephrin binding capability. **b,** Z-stack series of representative samples. Movies of the corresponding full set of fluorescent images are provided as online supplementary material. **c,** SF9 cells were transfected with recombinant baculovirus coding for mVenus-tagged EphA2 (green) or the mCherry-tagged ephrinA5$^{RBD}$ constructs (red). Cells expressing either protein do not cluster alone, but when mixed, they cluster due to EphA2-ephrinA5$^{RBD}$ interaction.




**Supplementary Fig. 7 | Interface mutations that affect EphA2 localization at the cell-cell contacts also affect EphA2 proteolytic cleavage.** mVenus tagged EphA2 point mutants were overexpressed in HEK293T cells in low fetal calf serum conditions (1.5% (v/v)). Cells were harvested in cold phosphate buffer saline, applied to a denaturing polyacrylamide gradient gel next to a poly-histidine-tagged protein ladder and proteins transferred to a nitrocellulose membrane. The part of the membrane corresponding to the ladder was immunoblotted with antibodies against poly-histidine, the rest of the membrane was immunoblotted with a mixture of anti-phosphotyrosine and anti-actin (top row) or a mixture of anti-GFP and anti-actin (bottom row). Additional Western blots using either anti-phosphotyrosine or anti-actin made the distinction between tyrosine phosphorylated bands and actin possible (not shown). As previously reported[4] for wild type EphA2, the immunoblot assays showed that full length EphA2 constructs, including the point mutants in interfaces A-F, were strongly tyrosine phosphorylated on over-expression in HEK293T cells. The arrows on the right indicate prominent bands: phosphorylated EphA2 (arrow 1), C-terminal EphA2 fragments that are phosphorylated (arrows 2-4), actin (arrow 5). Note that the fragments highlighted by arrows 2 and 3 are absent in mutants that lack specific localization to the cell-cell contacts (interface A, B, D mutants). The fragment shown by arrow 3 is present for Interface F mutants, but not that shown by arrow 2. The fragment shown by arrow 4 is present for all mutants.




**Supplementary Fig. 8 | Multiple angle light scattering (MALS).** MALS was used to determine that eEphA2 is monomeric in solution and that its complex with ephrinA5$^{RBD}$ has a 1:1 stoichiometry. **a,** Unliganded eEphA2, the refractive index is shown in red, expected mass/monomer = 57 kDa, measured mass = 59 kDa. **b,** eEphA2-ephrinA5$^{RBD}$ complex, the refractive index is shown in blue, expected mass for 1+1 chain complex = 75 kDa, measured mass = 73 kDa. The tail in the peak and slightly sloping measured mass may be due to co-migrating unliganded eEphA2. For all the measured masses the reported uncertainty was 4% as calculated by the ASTRA data analysis software.

## SUPPLEMENTARY METHODS

**Protein purification.** We expressed human EphA2 ectodomain (eEphA2) and human ephrin A5 ectodomain (ephrinA5$^{RBD}$) transiently[5] in GnTI-deficient HEK293S cells[6]. We dialyzed cell medium against phosphate buffer saline and loaded it onto a Ni-containing affinity column (His-Trap, GE Healthcare). We then washed the column with 20mM Tris pH 7.5, 300mM NaCl, 40mM imidazole and eluted the protein with 20mM Tris pH 7.5, 300mM NaCl, 230mM imidazole. For crystallization, we incubated eEphA2 at room temperature with 1:20 recombinant endoglycosidase F1[7-8] overnight. To remove the endoglycosidase, we diluted the protein 1:10 with 20mM Tris pH 7.5, 300mM NaCl, re-loaded it on a Ni-containing column, washed and re-eluted. We concentrated the eluted protein and loaded it on a Superdex 200 column (GE Healthcare) previously equilibrated with phosphate buffer saline (Sigma). Native protein failed to yield well diffracting crystals, we therefore methylated the lysine residues using an established protocol[9]. We added 100mM Tris pH 7.5 after the end of the methylation reaction. We concentrated methylated protein solution and loaded it onto a Superdex 200 column previously equilibrated with crystallization buffer (200mM NaCl and 20mM Tris pH 7.5). We pooled fractions from the main protein peak and concentrated these to 8.5 mg ml$^{-1}$ for crystallization. We carried out the expression of selenomethionine labeled protein according to previously published protocols[5]. We performed the purification and crystallization of selenomethionine derivative protein as for the native protein. To produce complex of eEphA2 and ephrinA5$^{RBD}$, we loaded media containing the two proteins separately on Ni-containing columns and eluted using the protocol described above. We then mixed eEphA2 and ephrinA5$^{RBD}$ in a 1:1.5 molar ratio, and subsequently deglycosylated, lysine-methylated and purified the complex on a Superdex 200 column using the protocol described above for eEphA2. We pooled the eluted fractions containing protein complex and concentrated these to 10 mg ml$^{-1}$.

**Structure refinement of eEphA2.** The higher resolution data set that we collected from a eEphA2 selenomethionine derivative crystal was anisotropic and extended up to a maximum resolution of 4-2.75 Å depending on the crystal orientation. We used all data up to the highest resolution (Friedel pairs merged) for model building as this improved the quality of electron density maps. In the highest resolution shell (2.85-2.75 Å) we measured 710 reflections (571 unique), thereby achieving 35% completeness, with I/σ(I) = 1.25, and Rmeas= 0.86. For the final rounds of refinement, and to conform to the currently accepted crystallographic procedure, we trimmed back the data 3.0 Å resolution, with an overall completeness of 98% (96% in the highest resolution shell 3.1-3.0 Å resolution). We have submitted all reflection data up to 2.75 Å resolution to the Protein Data Base (Friedel pairs not merged).

**Structure refinement of eEphA2-ephrinA5$^{RBD}$.** After molecular replacement, we performed rigid body refinement defining each domain as a separate rigid body (resulting R-work=0.368, R-free = 0.378), followed by two macro-cycles of TLS refinement (resulting Rwork=0.308, R-free = 0.326) and then regularized the geometry to remove any clashes (final R-work = 0.3123, R-free = 0.3143). The FN1-FN2 linker changes conformation compared to the unliganded molecular replacement model and linker

residues 433-437 were removed from the final complex model. Further, the loop comprising residues 513-515 (forming part of interface F) is not structured in the input model and was therefore also not included in the complex structure. The final electron density maps we obtained from the complex crystal show positive density for both these regions and the putative positions of the missing chains are indicated as dotted lines in **Supplementary Fig. 1**. Due to the low resolution of the data we did not attempt to build these regions *de novo* into the complex structure.

**Multiangle Light Scattering (MALS).** We purified proteins by size-exclusion chromatography and concentrated them to approximately 1.6 ml $ml^{-1}$ (eEphA2) and 1.8 mg $ml^{-1}$ (eEphA2/ephrinA5$^{RBD}$ complex). We achieved separation for the MALS using an analytical Superdex S75 10/30 column (GE Heathcare) and passed the eluate through online static light scattering (DAWN HELEOS II, Wyatt Technology, Santa Barbara, CA), differential refractive index (Optilab rEX, Wyatt Technology) and Agilent 1200 UV (Agilent Techologies) detectors. We analyzed data using the ASTRA software package (Wyatt Technology).

**Immunoblot analysis.** To assess cell surface expression of truncated EphA2 constructs, we incubated the cells for 10 minutes at 4 ˚C to slow down endocytosis and then for another 30 min with 1 mg $ml^{-1}$ No-Weigh Sulfo-NHS-LC-Biotin (Pierce) at 4 ˚C. We stopped the biotinylation reaction by adding 100 mM Tris pH 7.5. We lyzed the cells in phosphate buffer saline supplemented with 100 mM Tris pH 7.5 and 1% (v/v) Triton X-100 (Sigma). We incubated cell lysates for 30 minutes with streptavidin-coated agarose beads (Pierce) at 4 ˚C, washed the beads three times with lysis buffer and then re-suspended the beads in standard protein denaturing gel loading buffer for immunoblot analysis using anti-pentahistidine (penta-His, Qiagen, dilution 1:1000). We resuspended cells intended for anti-phosphotyrosine or anti-green fluorescent protein immunoblot analyses in chilled phosphate buffer saline 15-20 hours post transfection, immediately mixed them with standard protein gel loading buffer and analyzed the samples using anti-phosphotyrosine (4G10® Platinum, Millipore, dilution 1:2000), anti-GFP (Cat. No. A-11122, Invitrogen, dilution 1:1000), anti-actin (ACTN05 (C4), AbCam, dilution 1:400), anti-mouse IgG peroxidase-conjugated secondary antibody (Sigma, dilution 1:10000) and anti-rabbit IgG peroxidase-conjugated secondary antibody (AbCam, dilution 1:5000). To rule out that FCS or trace amounts of trypsin used for cell propagation are responsible for EphA2 cleavage, we performed control experiments in cells cultured with 0% or 1.5% (v/v) FCS, in the presence of 0%, 0.5% or 5% (v/v) of the trypsin-EDTA solution used for tissue culture maintenance (Invitrogen, Cat. No. 25300-054) (data not shown).

**Supplementary Table 1. Analysis of eEphA2-eEphA2 and eEphA2-ephrinA5$^{RBD}$ protein interfaces A-F.**

| Interface | Surface buried in Å$^2$* ; domains interacting in eEphA2-ephrinA5$^{RBD}$ complex crystals | Surface buried in Å$^2$* ; domains interacting in unliganded eEphA2 crystals | Site-directed EphA2 mutants studied | Proteolytic fragments like non-mutated EphA2 | Accumulation at cell-cell contacts like non-mutated EphA2 |
|---|---|---|---|---|---|
| A | 1983 LBD-RBD; major ephrin binding site | - | T101A+R103E | no | (no)[#] |
| | | | A190N + L192S | no | no |
| B | 1066 LBD- LBD | 1626 LBD-sushi domain | D129N + G131S | no | no |
| | | | G131Y | no | no |
| | | | T132E | no | no |
| C | 173 LBD-RBD | - | L138N | yes | yes |
| | | | L138Y | yes | yes |
| D | 972 sushi domain-sushi domain | 1626 LBD-sushi domain | H246N + A248S | no | no |
| | | | L254D+V255D | no | no |
| | | | L254D+V255D + I257D | no | no |
| E | 268 FN1domain-FN1 domain | - | P378E | yes | yes |
| | | | P378N + E380S | yes | yes |
| | | | E380G | yes | yes |
| F | 1623 FN1/FN2 domains- RBD | 175 FN2 domain-FN1 domain | E513R | no | (yes)[$] |
| | | | Deletion of residues 512-515 (Δ512-515) | no | (yes)[$] |

* PDBsum[47] result

[#] intermediate phenotype: mutant protein is mostly distributed homogenously on the entire cell surface, with weak accumulation at cell-cell contacts of transfected cells

[$] intermediate phenotype: mutant protein is observed on the entire cell surface, with strong accumulation at cell-cell contacts of transfected cells

## SUPPLEMENTARY REFERENCES


1 Bocharov, E. V. *et al.* Spatial structure and pH-dependent conformational diversity of dimeric transmembrane domain of the receptor tyrosine kinase EphA1. *J Biol Chem* **283**, 29385-29395 (2008).

2 Wybenga-Groot, L. E. *et al.* Structural basis for autoinhibition of the Ephb2 receptor tyrosine kinase by the unphosphorylated juxtamembrane region. *Cell* **106**, 745-757 (2001).

3 Stapleton, D., Balan, I., Pawson, T. & Sicheri, F. The crystal structure of an Eph receptor SAM domain reveals a mechanism for modular dimerization. *Nat Struct Biol* **6**, 44-49 (1999).

4 Miura, K., Nam, J. M., Kojima, C., Mochizuki, N. & Sabe, H. EphA2 engages Git1 to suppress Arf6 activity modulating epithelial cell-cell contacts. *Mol Biol Cell* **20**, 1949-1959 (2009).

5 Aricescu, A. R., Lu, W. & Jones, E. Y. A time- and cost-efficient system for high-level protein production in mammalian cells. *Acta Crystallogr D Biol Crystallogr* **62**, 1243-1250 (2006).

6 Reeves, P. J., Callewaert, N., Contreras, R. & Khorana, H. G. Structure and function in rhodopsin: high-level expression of rhodopsin with restricted and homogeneous N-glycosylation by a tetracycline-inducible N-acetylglucosaminyltransferase I-negative HEK293S stable mammalian cell line. *Proc Natl Acad Sci U S A* **99**, 13419-13424 (2002).

7 Grueninger-Leitch, F., D'Arcy, A., D'Arcy, B. & Chene, C. Deglycosylation of proteins for crystallization using recombinant fusion protein glycosidases. *Protein Sci* **5**, 2617-2622 (1996).

8 Chang, V. T. *et al.* Glycoprotein structural genomics: solving the glycosylation problem. *Structure* **15**, 267-273 (2007).

9 Walter, T. S. *et al.* Lysine methylation as a routine rescue strategy for protein crystallization. *Structure* **14**, 1617-1622 (2006).